Connection and Validation of Narrow-Band ΔVLBI Phase Observations

S. C. Wu

Tracking Systems and Applications Section

Two-station narrow-band ΔVLBI requires phase connections between consecutive scans. This article presents an efficient computer-aided scheme for this purpose. This scheme is an iteration process alternating between a grand fit on many scans and integer quantization of the phase-shift cycles to be assigned to the scans. Only linear simultaneous equations of a few unknowns need to be solved. A simulation analysis indicates that faultless phase connection can be expected when there is no localized systematic noise. When systematic noise of moderate level exists, the possible incorrect phase connection can be detected and corrected for by comparing the connected phases from the two alternating observations, after removing the residual diurnal effects.

I. Introduction

Narrow-band ΔVLBI (Differential Very Long Baseline Interferometry) has been considered as one of the possible accurate spacecraft tracking systems. A demonstration plan has been made for Voyager spacecraft during Jupiter encounters in 1979 (Ref. 1). As is well known, narrow-band ΔVLBI requires a long pass (3-4 hr) of tracking so that its diurnal signature can resolve, in two orthogonal directions, the angular separation between spacecraft (S/C) and an angularly nearby extragalactic radio source (EGRS). On the other hand, with 2-station mode, which is most likely the case, S/C and EGRS observations are taken alternately and gaps in each data stream are inevitable. The resulting interferometric phase of each continuous data segment (scan) has a $2n\pi$ ambiguity with different n for different segments. To recover the correct diurnal signature of each data stream, the interferometric phases of the corresponding scans have to be connected with correct $2n\pi$ phase shift assigned to each scan.

Such phase connection, if done manually by "eyeballing" the variations of the phases, is not only time consuming and tedious but may result in incorrect integer n since the phase change across the gap may be as large as tens of cycles.

This article presents an efficient computer-aided phase-connection scheme. A simulation analysis shows that the scheme works faultlessly when there is no systematic noise such as clock drift or localized media effects. When such localized fluctuation exists, the scheme may assign incorrect integer cycle adjustments to some scans. This can easily be detected and corrected for by a phase validation scheme.

It should be pointed out that for a successful phase connection, either manually or by a computer, the effect of the earth's spin should be modeled out and only the "residual phases" need be connected.

II. The Grand Fit

When manually connecting the phases, one would fit each scan by a straight line and then try to connect two consecutive scans at one time by shifting one of them up or down by $2n\pi$ such that they look continuous. This appears no problem only when the two fitted straight lines have the same slopes; ambiguity arises when the slopes are different, as may happen in practice. Higher-order fit may seem to resolve this problem. However, an independent higher-order fit on each scan is prone to error due to a limited number of noise-affected data points in a scan. Consequently, a grand fit on as many scans as possible simultaneously is desirable for a computer-aided phase-connection scheme.

It is well known that most, if not all, fitting (optimization) algorithms require the solved-for parameters to be continuous over the region of interest. However, the integers n of concern are discrete numbers. A grand fit with such constraint imposed[1] becomes nonlinear and annoying. The present scheme does the grand fit free from such constraint and allows the parameters n to assume any value. An iterative process is then performed to successively quantize these values of n into integers. The iterative process will be discussed in Section III.

Let the measured residual phase of the ith data point in the jth scan be denoted as $\phi_{i,j}$; let there be J scans to be included in a single grand fit and I_j data points in the jth scan. Then for a grand fit of order M one needs to minimize the following function with respect to a_m and x_j:

$$f = \sum_{j=1}^{J} \sum_{i=1}^{I_j} \left\{ \sum_{m=0}^{M} a_m t_{i,j}^m - (\phi_{i,j} + 2\pi x_j) \right\}^2 \qquad (1)$$

where $t_{i,j}$ is the time. The symbol x_j has been used in place of n to denote its role as a continuous variable.

Since only relative phases are needed we can arbitrarily set $x_1 = 0$. Then (1) is replaced by

$$f = \sum_{i=1}^{I_1} \left(\sum_{m=0}^{M} a_m t_{i,1}^m - \phi_{i,1} \right)^2$$
$$+ \sum_{j=2}^{J} \sum_{i=1}^{I_j} \left(\sum_{m=0}^{M} a_m t_{i,j}^m - \phi_{i,j} - 2\pi x_i \right)^2 \qquad (2)$$

[1] For instance, a penalty function $(1 - \cos^{2m} \pi x)$ can be added onto the function to be minimized (cf. Eq. 2).

To minimize (2) its partial derivatives with respect to a_0, $a_1, \ldots, a_M$ and $x_2, x_3, \ldots, x_j$ are set to zero. With some manipulations these become

$$\sum_{m=0}^{M} \sum_{j=1}^{J} \left(\sum_{i=1}^{I_j} t_{i,j}^{m+r} \right) a_m - 2\pi \sum_{j=2}^{J} \left(\sum_{i=1}^{I_j} t_{i,j}^r \right) x_j =$$

$$\sum_{j=1}^{J} \sum_{i=1}^{I_j} t_{i,j}^r \phi_{i,j} \qquad (3)$$

$$r = 0, 1, \ldots, M$$

and

$$\sum_{m=0}^{M} \left(\sum_{i=1}^{I_j} t_{i,j}^m \right) a_m - (2\pi I_j) x_j = \sum_{i=1}^{I_j} \phi_{i,j}, \quad j = 2, 3, \ldots, J \qquad (4)$$

Equations (3) and (4) are $M + J$ simultaneous linear equations for the $M + J$ unknowns a_m and x_j. A further reduction of this system of linear equations can be done by substituting each x_j from (4) into (3). This yields

$$\sum_{m=0}^{M} \left\{ \sum_{i=1}^{I_1} t_{i,1}^{m+r} + \sum_{j=2}^{J} \left[\sum_{i=1}^{I_j} t_{i,j}^{m+r} - \left(\sum_{i=1}^{I_j} t_{i,j}^m \right) \left(\sum_{i=1}^{I_j} t_{i,j}^r \right) \Big/ I_j \right] \right\} a_m = \sum_{i=1}^{I_1} t_{i,1}^r \phi_{i,1} + \sum_{j=2}^{J} \left[\sum_{i=1}^{I_j} t_{i,j}^r \phi_{i,j} - \left(\sum_{i=1}^{I_j} t_{i,j}^r \right) \left(\sum_{i=1}^{I_j} \phi_{i,j} \right) \Big/ I_j \right] \qquad (5)$$

$$r = 0, 1, \ldots, M$$

Hence, a system of $M + 1$ linear equations for a_m is resulted. This can be written in the following matrix form:

$$[s] \{a\} = \{b\} \qquad (6)$$

where [ ] denotes a square matrix and { } a column matrix. The elements of $[s]$ and $\{b\}$ are

$$S_{r,m} = S_{m,r} =$$

$$\sum_{i=1}^{I_1} t_{i,1}^{m+r} + \sum_{j=2}^{J}\left[\sum_{i=1}^{I_j} t_{i,j}^{m+r} - \left(\sum_{i=1}^{I_j} t_{i,j}^{m}\right)\left(\sum_{i=1}^{I_j} t_{i,j}^{r}\right)\Big/I_j\right]$$

(7)

and

$$b_r = \sum_{i=1}^{I_1} t_{i,1}^{r}\,\phi_{i,1} + \sum_{j=2}^{J}\left[\sum_{i=1}^{I_j} t_{i,j}^{r}\,\phi_{i,j}\right.$$

$$\left. - \left(\sum_{i=1}^{I_j} t_{i,j}^{r}\right)\left(\sum_{i=1}^{I_j} \phi_{i,j}\right)\Big/I_j\right]$$

(8)

In practice, $M = 2$ or 3 will suffice and the solutions of a_m are straightforward. Once the coefficients a_m have been determined, the phase shift parameters x_j are calculated according to (4):

$$x_j = \frac{1}{2\pi I_j}\left[\sum_{m=0}^{M}\left(\sum_{i=1}^{I_j} t_{i,j}^{m}\right)a_m - \sum_{i=1}^{I_j}\phi_{i,j}\right], \quad j = 2, 3, \ldots, J.$$

(9)

III. Iteration Process of Integer Quantization for x_j

The phase-shift parameters x_j determined by a grand fit in the preceding section will in general differ from integers. If the differences are all small, one may conceivably set them equal to the integers they approximate. However, if the differences are not all small, one may hesitate to equate x_j to an integer. However, the following iteration bears out the method.

From (1) it is obvious that any variation in x_j from correct integers induces corresponding changes in the fitted coefficients a_m. Let the change in each a_m be σ_m, $m = 0, 1, \ldots, M$. Then the variation in x_j can be expressed, according to (9), as

$$\sigma_j = \frac{1}{2\pi I_j}\sum_{m=0}^{M}\left(\sum_{i=1}^{I_j} t_{i,j}^{m}\right)\sigma_m$$

(10)

This implies that the variation x_j from the correct integers increases with $t_{i,j}$. In other words, the first few x_j are much closer to the correct integers than the remaining x_j. Hence, one can comfortably set the first few x_j to the nearest integers. With these integers fixed, another grand fit is performed to improve a_m and the remaining x_j. The first few improved x_j will now be closer to the correct integers and can be quantized with greater confidence. This process is repeated until all x_j are quantized into integers.

After each grand fit, the criterion of setting the first x_j to the nearby integer can be more relaxed, as this x_j is more likely to be very nearly an integer. On the other hand it is more probable for the remaining x_j's to miss the correct integer values by an amount greater than 1/2 and the criterion should be made more stringent. In the proposed scheme the allowable variation is chosen to be 0.4 for the first x_j after each grand fit; for each of the remaining x_j, both x_j and x_{j-1} must have a variation < 0.1.

IV. Simulation Analysis

The flow chart of the phase-connection scheme is shown in Fig. 1. The order of the polymonial to be fitted is assigned. In most cases a second- or third-order polymonial will be sufficient. The simulated phase data are generated with the following parameters:

data density = 1 per minute

gap width = W = an integer multiple of a minute

number of data points per scan = I, a variable parameter

data noise = σ, a variable parameter

pass length = 3 hours

number of scans = $J = (3 \times 60)/(W + I)$ rounded to the next lowest integer

$a_1 = 0.5$ rad/s

$a_2 = 0, -5 \times 10^{-5}, -1 \times 10^{-4}$ rad/s^2

$a_3 = 0, 5 \times 10^{-9}$ rad/s^3

$$\phi_{i,j} = \sum_{m=1}^{3} a_m\, t_{i,j}^{m} + \sigma - 2n_j\pi \text{ with } n_j \text{ to be such that}$$

$$0 \leqslant \phi_{1,j} < 2\pi.$$

Both $M = 2$ and $M = 3$ are tried for the fit. The following inferences are drawn from the simulation:

(1) While higher order M for the polynomial to be fitted may be chosen even when $a_2 = a_3 = 0$ in the simulated

phase data, a lower-order polynomial is less susceptible to data noise. Hence one may start with a higher-order polynomial and decrease the order upon failure in phase connection.

(2) As few as 2 data points per scan are allowable for a successful phase connection provided the gap width and the data noise are reasonably small.

(3) Given a data density and a gap-width-to-data-span ratio the success in phase connection degrades as data noise σ increases but seems independent of the gap width, at least for gap width $\leqslant 10$ minutes. This relaxes the gap width limit for minimum data loss when time offset between S/C and EGRS observations is called for (Ref. 2).

(4) The maximum data noise above which phase connection may fail decreases with increasing gap-width-to-data-span ratio. Figure 2 is an example with 5-minute data span (per scan) over three hours. When there is no localized fluctuation due to systematic noise, a typical ΔVLBI pass[2] will fall well within the convergence range, as shown in the figure.

V. Grouping of Scans in a Single Grand Fit

As pointed out in Section II, it is desirable to include as many scans as possible in a single grand fit. However, a problem arises when sizable localized phase fluctuations exist among scans included in a single grand fit: A low-order polynomial may lose track of these fluctuations while a polynomial of higher order may degrade the convergence of the iteration process. Hence an appropriate grouping of scans in a single fit is essential.

A criterion to determine the grouping of scans is the change in slope of the straight lines fitted to the consecutive scans. Two consecutive scans are to be grouped together if the slopes of their fitted straight lines satisfy

$$|S_j - S_{j-1}| - (\sigma_j^2 + \sigma_{j-1}^2)^{1/2} \leqslant 1 \text{ mHz} \qquad (11)$$

where S_j and σ_j are the slope and its uncertainty of scan j. The number of scans J in a group increases until (11) fails. When this number of scans J is determined, the order of polynomial is selected to be

$$M = J, \quad \text{if} \quad J \leqslant 5$$
$$M = 5, \quad \text{if} \quad J > 5 \qquad (12)$$

The order of fit M is to be decreased by 1 upon failure of convergence in the iteration and a new fit is tried. If the iteration does not converge for all $M \geqslant 2$ the last scan in the group is removed from the fit, the largest M according to (12) is selected and the process repeats.

It is obvious that phase connection is redundant for a group with fewer than two scans. When the number of scans reduces to 2, due to the failure of either (11) or in the iteration process, the slope test of (11) is to be ignored. Also, the allowable deviation of the fitted x_j away from an integer is to be relaxed to 0.5 to exclude any failure in the iteration process. The actual deviation ϵ_{x_j} of each fitted x_j from its quantized value can serve as an indication of the degree of confidence in the resulting phase connection.

VI. Validation of Connected Phases

The phase-connection scheme discussed in the preceding sections works faultlessly when there is no localized phase fluctuations. Because of irregularities in transmission media (ionosphere in particular) and instability of electronic signal path, localized fluctuations do exist. When such fluctuations are sizable, not only do we need to divide the scans into several groups in connecting phases, but also some of the scans may be assigned incorrect cycle adjustments. To guard against such danger, the following phase validation process will be necessary.

Since the residual phases from S/C and EGRS are each independently connected, an approach of phase validation is to bring the two connected phases together and examine whether their variation is continuous over the pass. However, such examination is indicative only if the two residual phases have identical residual model errors. The residual model errors in frequency offset are identical since common frequency standards are used for both signals. Baseline error is not identical even though the same baseline is used, owing to different baseline projections onto the planes of sky in the two different directions of the two sources. Angular position (right ascension and declination) error is also different. Since both baseline error and angular position error appear in the form of residual diurnal variation, they can be easily removed by fitting $A \cos \omega_e t + B \sin \omega_e t$ to each of S/C and EGRS residual phases, where ω_e is the earth's spinning rate. In practice, two more terms, $Ct + D$, are needed to remove the residual frequency offset and phase offset. This is a simple linear least-squares fit with four degrees of freedom.[3]

[2]S/N = 5 corresponding to $\sigma = 0.2$ rad, and gap width/data span = 1.2.

[3]Alternatively, a simultaneous fit for S/C and EGRS phases with common frequency offset term can be adopted. The degrees of freedom become 7.

After removing the residual model errors, the S/C and EGRS phases are plotted on the same graph. This can be done handily by print-plotting on standard computer print. Now comes the time for inspection. Unfortunately, this has to be done by eyeballing at the present time. This is best illustrated by an example of the actual ΔVLBI pass. Figure 3 shows the connected phases of Voyager 1 and OJ 287 with their residual model errors removed. Two consecutive scans from the same source (S/C or quasar) are considered "continuous" if the phase change across the gap is consistent with the phase change in the intervening scan from the other source. A quick glance over the graph reveals that there are discontinuities between scans 2 and 3 of S/C phase and between the last two scans of quasar phase. When these two 1-cycle adjustments are made, the phases after removing the updated residual model errors are shown in Fig. 4. No obvious discontinuity remains and the adjustments are validated. Normally only one or two iterations of adjustment are needed to eliminate all offensive discontinuities.

There has been concern that when a clock without extremely high stability is used, phase connection may fail. By the above phase validation scheme, incorrectly connected phases can still be detected and corrected as long as the phase drift due to clock instability is continuous (i.e., no phase jumps occur). Figure 5 is an example where rubidium clock was used at one end of the baseline. A discontinuity is detected between the last second and third scans of the S/C phase. After a 1-cycle adjustment the phases are shown in Fig. 6. Again, no obvious discontinuity remains. It should be noted that the gross effect of the clock drift has been absorbed by the fitting functions and only the residual "random walk" effect remains.

VII. Conclusions

A computer-aided, phase-connection scheme has been developed. A simulation analysis shows that, when there is no sizable localized phase fluctuation, this scheme connects VLBI residual phases faultlessly. When there are localized phase fluctuations the scheme may assign incorrect cycle adjustments to some of the scans. This can easily be detected and corrected for by a phase validation scheme. Currently this validation process requires an eyeballing inspection. A means to validate the connected phases automatically is currently being sought.

References

1. Brunn, D. L., et at., DSN Progress Report, 42-45, 15 June 1978, pp. 111-132.

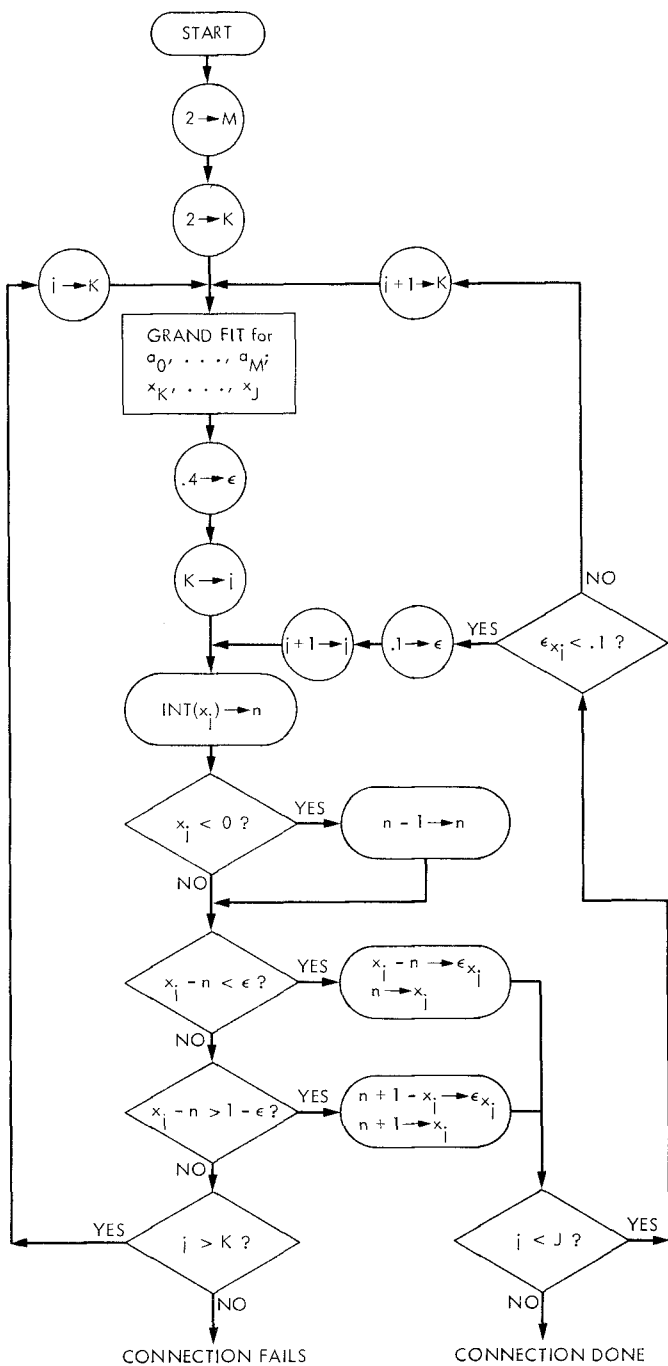2. Wu, S. C., IOM 315.1-296, June 12, 1978 (JPL internal document).

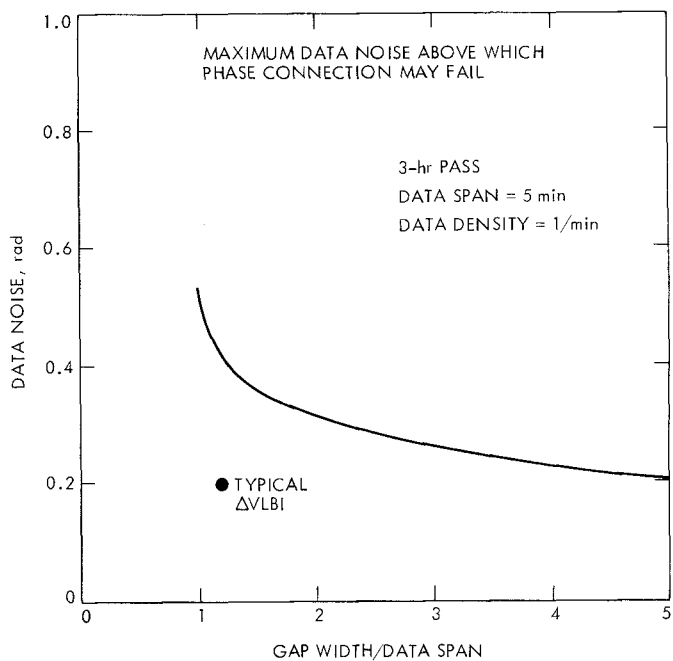Fig. 1. Flow chart of phase-connection scheme

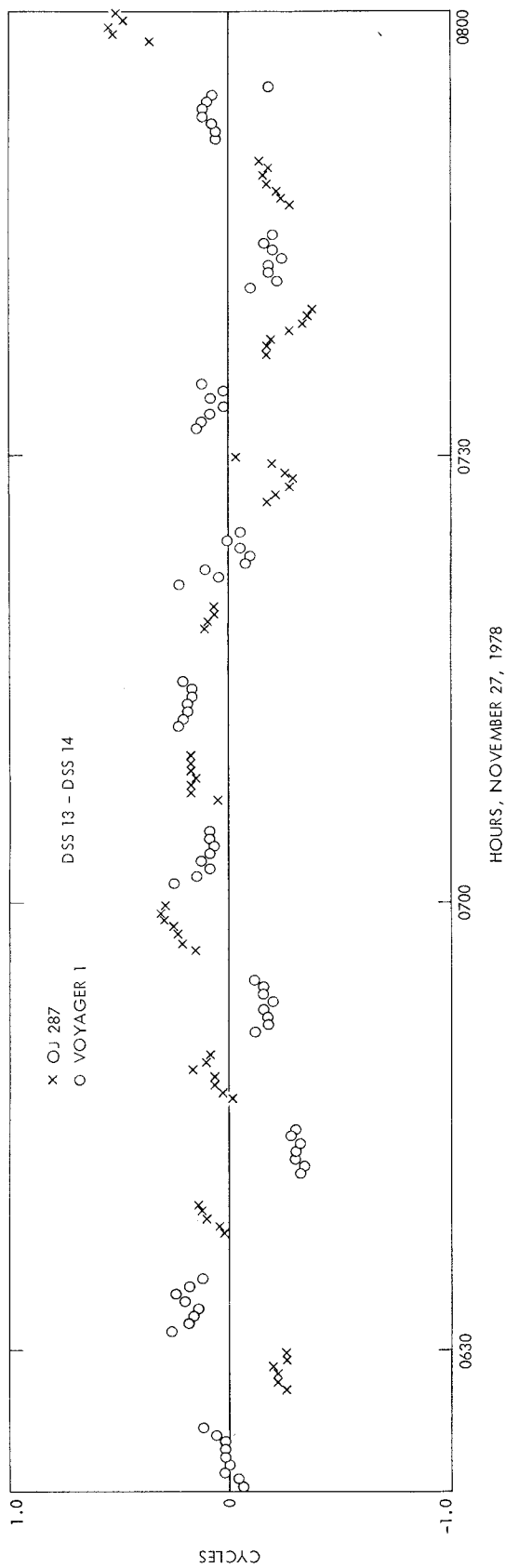Fig. 2. Phase-connection threshold for a scheme in Fig. 1

Fig. 3. Connected phases after removing constant rate and diurnal effects
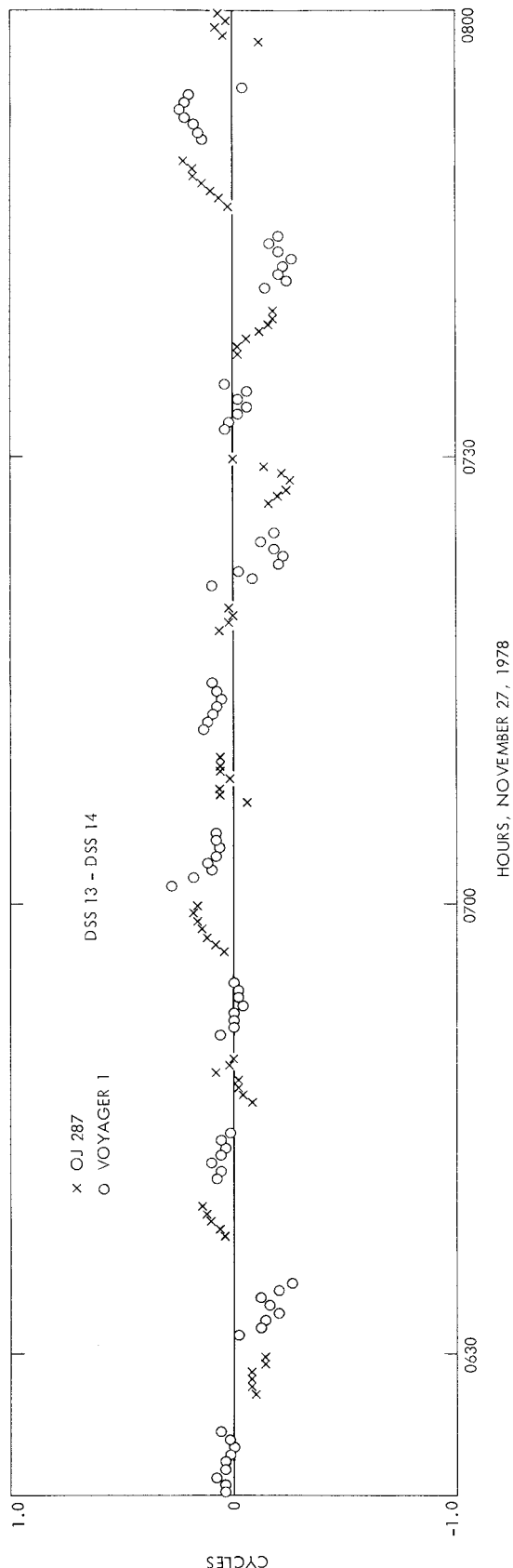
Fig. 4. Same as Fig. 3 except with cycle corrections
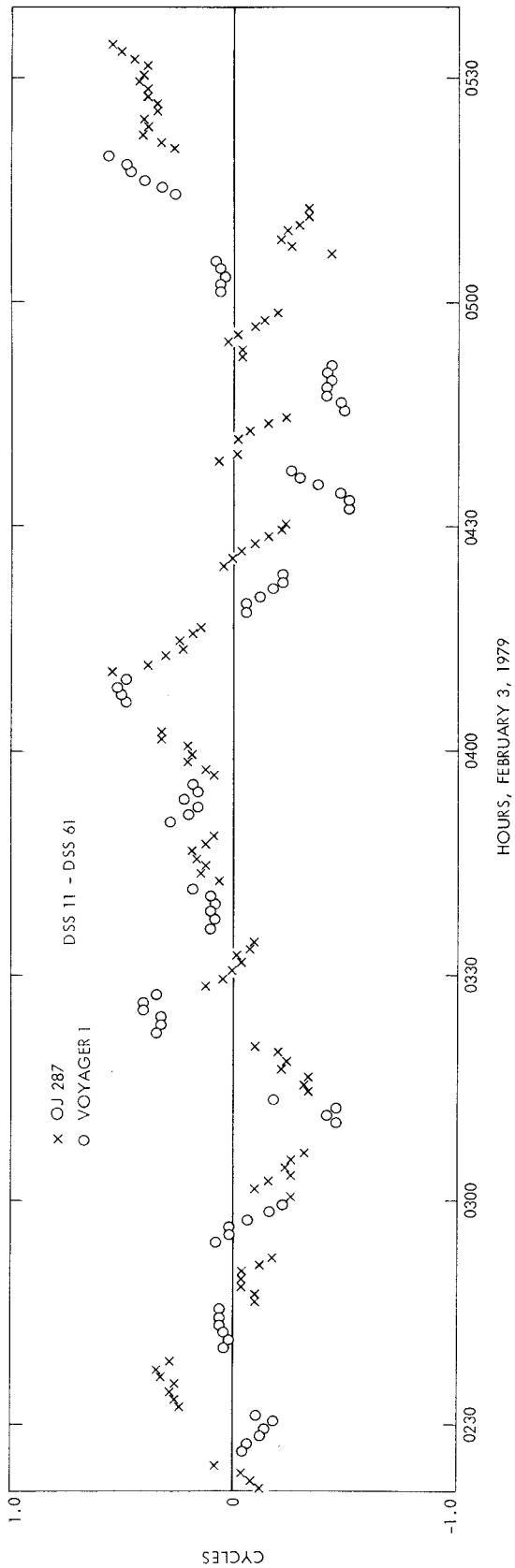
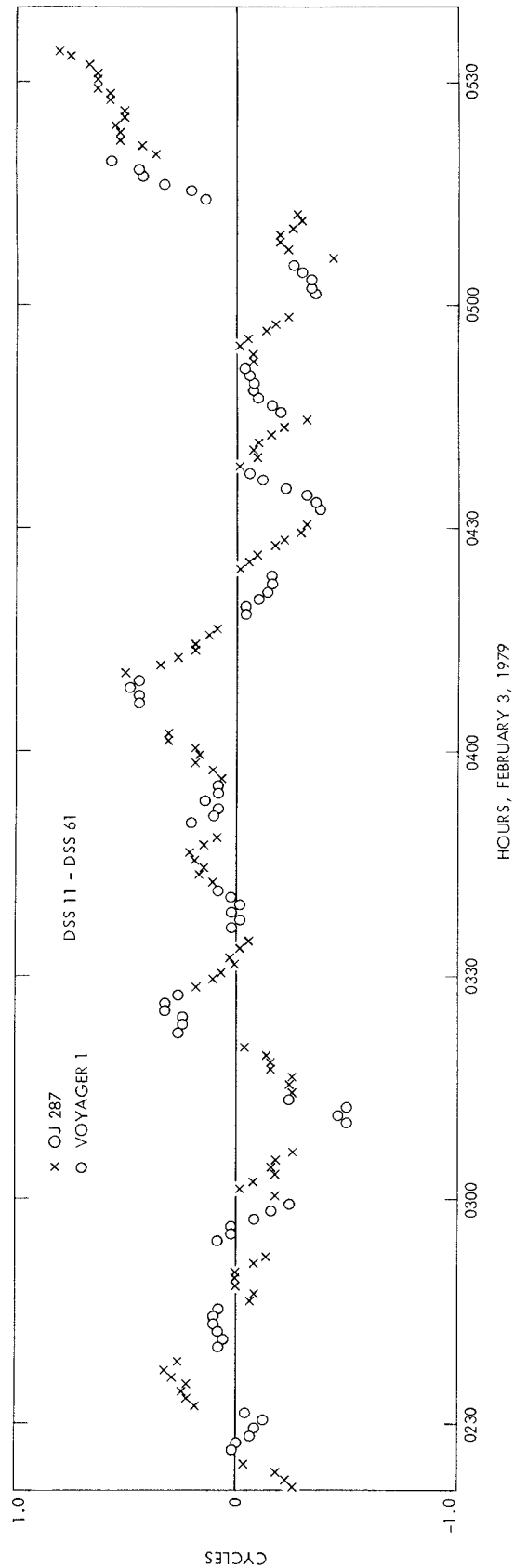Fig. 5. Connected phases after removing constant phase rate and diurnal effects

Fig. 6. Same as Fig. 3 except with cycle corrections

20